ON THE CONCENTRATION OF THE DOMINATION NUMBER OF THE RANDOM GRAPH

ROMAN GLEBOV, ANITA LIEBENAU, AND TIBOR SZABÓ

ABSTRACT. Many natural graph parameters of the Erdős-Rényi random graph $\mathcal{G}(n,p)$ are concentrated on two values, provided the edge probability p is sufficiently large. In this paper we study the domination number of $\mathcal{G}(n,p)$. Wieland and Godbole proved that the domination number of $\mathcal{G}(n,p)$ is equal to one of two values asymptotically almost surely whenever p is constant or just slightly sub-constant. We extend this result for the range when p is of larger order than $\frac{\ln^2 n}{\sqrt{n}}$. We also indicate why some sort of lower bound on the probability p is necessary. Concentration, though not on a constant length interval is also proven for p as small as $\omega(1/n)$. For the range $p = \Theta(1/n)$, we show under a plausible assumption the length of the smallest interval on which the domination number is a.a.s. concentrated.

1. INTRODUCTION

In this paper we investigate the domination number of the homogeneous random graph model $\mathcal{G}(n,p)$. As usual, a graph $G \sim \mathcal{G}(n,p)$ is a graph with vertex set $[n] = \{1,\dots,n\}$ in which edges are chosen to be inserted independently with probability p. In a graph $G = (V,E)$, we call a set $S \subseteq V$ *dominating* if every vertex $v \in V$ is either a member of S, or adjacent to a member of S. The *domination number $D(G)$* is the smallest cardinality of a dominating set in G. Early results on the concentration of the domination number of $G \sim \mathcal{G}(n,p)$ include the case when p is fixed (see for example [?] and [?]), or when p tends to 0 sufficiently slowly. In this direction, Wieland and Godbole [?] showed that under the condition that either p is constant, or p tends to 0 with

$$p = p(n) \geq 10\sqrt{\frac{\ln\ln n}{\ln n}},$$

the domination number $D(G)$ takes one of two consecutive integer values with probability tending to 1, as n tends to infinity. In [?] it is raised as an open problem whether the validity of this 2-point concentration result can be extended to a wider range of p. In our main theorem we extend this range down to $p \gg \frac{\ln^2 n}{\sqrt{n}}$ and also include the range when $p \to 1$.

Here and in the rest of the paper, we denote $q = \frac{1}{1-p}$ and $d = np$. A statement about $G \sim \mathcal{G}(n,p)$ is said to hold *asymptotically almost surely (a.a.s.)* if it holds with probability tending to 1 as $n \to \infty$.

Date: November 28, 2012.

The first author was supported by DFG within the research training group "Methods for Discrete Structures".

The second author was supported by DFG within the graduate school Berlin Mathematical School.

Theorem 1.1. *Let $p = p(n)$ be such that $\frac{\ln^2 n}{\sqrt{n}} \ll p < 1$, and let $G \sim \mathcal{G}(n, p)$. Then there exists an $\hat{r} = \hat{r}(n, p(n))$, which is of the form*

$$\hat{r}(n, p(n)) = \log_q \left(\frac{n \ln q}{\ln^2 d} (1 + o(1)) \right),$$

such that $D(G) = \lfloor \hat{r} \rfloor + 1$ or $D(G) = \lfloor \hat{r} \rfloor + 2$ a.a.s.

Note that, when $p \to 0$ and hence $\ln q = p(1 + o(1))$ then the $\hat{r}$ from Theorem 1.1 is of the order $\frac{\ln d}{p} = \frac{n \ln d}{d}$.

The choice of the value of $\hat{r}$ in Theorem 1.1 as the start of the concentration interval will be a very natural one: $\hat{r}$ will represent a particular critical dominating set size, such that the expected number of dominating sets of size $\lfloor \hat{r} \rfloor$ tends to 0, whereas the expected number of dominating sets of size $\lfloor \hat{r} \rfloor + 2$ tends to ∞ very fast. For technical reasons the value of $\hat{r}$ will be defined somewhat implicitly in the $p \to 0$ range:

$$\hat{r} = \min \left\{ r \mid \mathbb{E}(X_r) \geq \frac{1}{d} \right\} - 1. \tag{1}$$

For $p \to 1$ (and $p < 1$), we will show that the theorem holds with the explicit formula $\hat{r} = \log_q \left(\frac{n \ln q}{\ln^2 n} \right)$.

Similarly to [?] we use standard first and second moment methods to prove the 2-point-concentration result of Theorem 1.1, however the technical difficulties increase significantly. A lower bound of $n^{-1/2}\mathrm{polylog}\, n$ on p seems to be the boundary of these calculations.

One can nevertheless show some, though not constant-length, concentration of $D(\mathcal{G}(n, p))$ for smaller p as well.

Proposition 1.2. *For $p \to 0$ and $d \to \infty$ we have that*

$$\mathbb{P} \left(D \left(\mathcal{G}(n, p) \right) = n \frac{\ln d}{d} (1 + o(1)) \right) \to 1.$$

If p tends to 0 faster than $n^{-1/2} \ln n$, then one can decrease the length of the concentration interval using Talagrand's Inequality.

Theorem 1.3. *Let $m = m(n)$ be the median of $D(G)$ where $G \sim \mathcal{G}(n, p)$ and let $t = t(n) \gg \sqrt{n}$. Then $\mathbb{P}(|D(G) - m| > t) \to 0$.*

By Proposition 1.2 $m \approx n \frac{\ln d}{d}$ so Theorem 1.3 constitutes an improved concentration result whenever $p \ll \frac{\ln n}{\sqrt{n}}$. However, for larger p, the length $t \gg \sqrt{n}$ of the interval of concentration is of larger order than the median.

In Theorem 1.1 we show that for $p \gg \ln^2 n / \sqrt{n}$ the domination number of $\mathcal{G}(n, p)$ is concentrated on two integers $\lfloor \hat{r} \rfloor + 1$ and $\lfloor \hat{r} \rfloor + 2$ a.a.s., where $\hat{r}$ is the size when the expected number of dominating sets changes from tending to zero to tending to infinity. It is natural to ask whether the lower bound, or at least some sort of lower bound, on p in Theorem 1.1 is justified. It is not hard to see that the theorem cannot be extended to hold for arbitrary p. For $p \ll n^{-4/3}$, for example, we have that $\mathcal{G}(n, p)$ consists of a collection of vertex-disjoint stars (with at most two edges) a.a.s., and hence its domination number equals n minus its number of edges. In Section 6 we extend this argument and show that for every $p = o(1/n)$, the domination number is not a.a.s. concentrated on any interval of length $o\left(\sqrt{\binom{n}{2} p}\right)$ a.a.s.

In our next theorem, we prove a somewhat weaker non-concentration result for larger p. We show that the domination number of $\mathcal{G}(n,p)$ is not concentrated around the critical $\hat{r}$.

Theorem 1.4. *For every $c, \varepsilon > 0$, there exists $\delta > 0$ such that for every $p = p(n)$ with $\varepsilon/n \leq p \ll 1$, there exists n_0 such that for every $n \geq n_0$, $D(\mathcal{G}(n,p)) > \hat{r} + c\frac{\hat{r}}{n\sqrt{p}}$ with probability at least δ.*

Observe that Theorem 1.4 implies that when $p = \Omega(1/n)$ and $p \ll (\ln n/n)^{2/3}$, the domination number $D(\mathcal{G}(n,p))$ is not concentrated on a constant length interval around $\hat{r}$, since $\frac{\hat{r}}{n\sqrt{p}} \sim \frac{\ln d}{np^{3/2}} \gg 1$. Furthermore, recall that Theorem 1.3 implies that for every $t \gg \sqrt{n}$, there exists an interval $I \subset [n]$ of length t such that $D(\mathcal{G}(n,p)) \in I$ a.a.s. For $p = \Theta(1/n)$, Theorem 1.4 states for every $c \in \mathbb{R}$ and sufficiently large n that $D(\mathcal{G}(n,p)) > \hat{r} + c\sqrt{n}$ with positive constant probability.

Notation and structure of the paper.
Throughout the paper, $\ln n$ denotes the natural logarithm of n.
Let $G \sim \mathcal{G}(n,p)$ and $r \leq n$. In Section 2, we examine the expected number of dominating sets of size r in G and define the critical $\hat{r}$ of Theorem 1.1 and 1.4 precisely. In Section 3 we prove Theorem 1.1. We split the proof according to whether $p \to 0$ or $p \to 1$. Section 4 contains the proofs of Proposition 1.2 and Theorem 1.3. In Section 5 we prove Theorem 1.4. Finally, in Section 6 we discuss some possible extensions of these results.

2. Expectation

In this section we study the expected number of dominating sets of size r in the random graph $\mathcal{G}(n,p)$ when $p \to 0$ and when $p \to 1$. In particular, we are interested in the value of r when the expectation first exceeds 1 and how fast it grows around this point. Let $G \sim \mathcal{G}(n,p)$. We denote by X_r the number of dominating sets of size r in G. For any fixed r-subset S of $[n]$ and vertex $v \in [n] \setminus S$ the probability that v is not dominated by S in G is $(1-p)^r$. These events are mutually independent for a fixed S, hence the probability that S is dominating is $(1-(1-p)^r)^{n-r}$ and for the expectation of X_r we have

$$\mathbb{E}(X_r) = \binom{n}{r}(1-(1-p)^r)^{n-r}.$$

It turns out that $\mathbb{E}(X_r)$ first exceeds 1 when r is in the range of $\log_q\left(\frac{n\ln q}{\ln^2 d}(1+o(1))\right)$. Recall that we use the notation $d = np$ and $q = \frac{1}{1-p}$. We now split the analysis according to whether $p \to 0$ or $p \to 1$.

2.1. The sparse case.

Note that when $p \to 0$, $\ln q = p(1+o(1))$. We thus have the identity $\log_q\left(\frac{n\ln q}{\ln^2 d}(1+o(1))\right) = \log_q\left(\frac{d}{\ln^2 d}(1+o(1))\right)$. The following two small calculations will come in handy.

Observation 2.1. *Let $p = p(n) \to 0$, $d = pn \to \infty$, and $r = \log_q\left(\frac{d}{\ln^2 d}(1+o(1))\right)$. Then the following identities hold.*

(i) $r = \frac{\ln d - 2\ln\ln d + o(1)}{\ln q} = \frac{\ln d}{p}(1+o(1)) = \frac{n\ln d}{d}(1+o(1))$.

(ii) $(1-p)^r = \frac{\ln^2 d}{d}(1+o(1)) \to 0$.

In the next lemma we establish that when r is in the range of our interest, then the expected number of dominating sets of size $r+1$ is much more than the expected number of dominating sets of size r.

Lemma 2.2. *Let $p = p(n) \to 0$, $d = pn \to \infty$. For $\ell < n/2$, $\mathbb{E}(X_l) < \mathbb{E}(X_{l+1})$. Furthermore, for every $r = r(n) = \log_q\left(\frac{d}{\ln^2 d}\right) + o(1/p)$ and $\alpha = o(1/p)$ we have*

$$\frac{\mathbb{E}(X_{r+\alpha})}{\mathbb{E}(X_r)} = \exp\Big((1+o(1))\,\alpha\,\ln^2 d\Big).$$

Remark 2.3. Note that $o(1/p) = \log_q(1+o(1))$ when $p \to 0$. Hence, the error terms in Observation 2.1 and Lemma 2.2 are of the same order.

Proof. First, let us note that, since $\alpha = o(1/p)$,

$$(1-p)^\alpha = 1 - p\alpha(1+o(1)). \tag{2}$$

For the lower bound, note that for every ℓ, $0 \le \ell < n/2$,

$$\begin{aligned}
\frac{\mathbb{E}(X_{\ell+1})}{\mathbb{E}(X_\ell)} &= \frac{\binom{n}{\ell+1}(1-(1-p)^{\ell+1})^{n-\ell-1}}{\binom{n}{\ell}(1-(1-p)^\ell)^{n-\ell}}\\
&= \frac{n-\ell}{\ell+1}\exp\left[(n-\ell)\sum_{k=1}^{\infty}\frac{(1-p)^{\ell k}}{k}\left(1-(1-p)^k\left(1-\frac{1}{n-\ell}\right)\right)\right]\\
&> \exp\left[(n-\ell)(1-p)^\ell\left(1-(1-p)\left(1-\frac{1}{n-\ell}\right)\right)\right]\\
&> \exp\left[(n-\ell)(1-p)^\ell p\right].
\end{aligned}$$

In particular, $\mathbb{E}(X_l) < \mathbb{E}(X_{l+1})$ for all $\ell < n/2$. Using the above, we estimate the quotient from the lemma as follows:

$$\begin{aligned}
\frac{\mathbb{E}(X_{r+\alpha})}{\mathbb{E}(X_r)} &= \frac{\mathbb{E}(X_{r+\alpha})}{\mathbb{E}(X_{r+\alpha-1})}\cdot\frac{\mathbb{E}(X_{r+\alpha-1})}{\mathbb{E}(X_{r+\alpha-2})}\cdot\dots\cdot\frac{\mathbb{E}(X_{r+1})}{\mathbb{E}(X_r)}\\
&\ge \exp\left[\sum_{i=0}^{\alpha-1}(n-(r+i))(1-p)^{r+i}p\right]\\
&\ge \exp\left[(n-r-\alpha)(1-p)^r\frac{1-(1-p)^\alpha}{p}p\right]\\
&= \exp\left[(1+o(1))\alpha\,\ln^2 d\right],
\end{aligned}$$

where in the last equality we used that $n-r-\alpha = n(1+o(1))$ by Observation 2.1(i), as well as (2) and the formula of Observation 2.1(ii).

For the upper bound note that for every $\ell \ge 0$, we have

$$\frac{\mathbb{E}(X_{\ell+1})}{\mathbb{E}(X_\ell)} = \frac{\binom{n}{\ell+1}(1-(1-p)^{\ell+1})^{n-\ell-1}}{\binom{n}{\ell}(1-(1-p)^\ell)^{n-\ell}} = \frac{n-\ell}{\ell+1}\left(1+\frac{p(1-p)^\ell}{1-(1-p)^\ell}\right)^{n-\ell}\cdot\frac{1}{1-(1-p)^{\ell+1}}$$

To estimate from above, we use again the telescopic product, and that by Observation 2.1 (*ii*) the last factor of the above expression is less than 2 for large n. Thus we obtain

$$\begin{aligned}
\frac{\mathbb{E}(X_{r+\alpha})}{\mathbb{E}(X_r)} &\leq \left(\frac{n}{r}\right)^{\alpha} \cdot \exp\left[\sum_{i=0}^{\alpha-1}(n-(r+i))\frac{(1-p)^{r+i}p}{1-(1-p)^{r+i}}\right] \cdot 2^{\alpha} \\
&\leq \left(\frac{2n}{r}\right)^{\alpha} \cdot \exp\left[\frac{n(1-p)^r p}{1-(1-p)^r} \cdot \sum_{i=0}^{\alpha-1}(1-p)^i\right] \\
&= \left(\frac{2n}{r}\right)^{\alpha} \cdot \exp\left[(1+o(1))np(1-p)^r \frac{1-(1-p)^{\alpha}}{p}\right] \\
&\leq \exp\left[\alpha\left(\ln(2n/r)+(1+o(1))np(1-p)^r\right)\right] \\
&= \exp\left[(1+o(1))\alpha \ln^2 d\right],
\end{aligned}$$

where in the last inequality we again use (2), and the last equality holds since by Observation 2.1 (*ii*)

$$\ln(2n/r) = \ln\left((2+o(1))d/\ln d\right) = o\left(\ln^2 d\right).$$

□

By the previous lemma, in our range of interest the expectation $\mathbb{E}(X_r)$ is strictly increasing in r and grows by a factor of $\exp\left[(1+o(1))\ln^2 d\right]$ with each increase of r by 1. Recall that our critical dominating set size is

$$\hat{r} = \min\left\{r \mid \mathbb{E}(X_r) \geq \exp\left[-\ln d\right]\right\} - 1.$$

Lemma 2.4. *Let $p \to 0$ and $d = np \to \infty$. Then $\hat{r}$ is of the form $\hat{r} = \log_q\left(\frac{d}{\ln^2 d}(1+o(1))\right)$. Furthermore*

(i) $\mathbb{E}(X_{\hat{r}}) \to 0$, *and*
(ii) $\mathbb{E}(X_{\hat{r}+2}) \geq \exp\left[(1+o(1))\ln^2 d\right] \to \infty$.

Proof. First let $r = \left\lfloor \log_q\left(\frac{d}{\ln^2 d}\right)\right\rfloor$. Then Observation 2.1 applies, so that

$$\begin{aligned}
\mathbb{E}(X_r) &= \binom{n}{r}(1-(1-p)^r)^{n-r} \\
&\leq \exp\left[r\ln\left(\frac{ne}{r}\right) - (n-r)(1-p)^r\right] \\
&\leq \exp\Bigg[\left(\frac{\ln d}{\ln q} - \frac{2\ln\ln d}{\ln q}(1+o(1))\right)\cdot\ln\left(\frac{de}{\ln d}(1+o(1))\right) \\
&\qquad\qquad - \frac{n\ln^2 d}{d} + \frac{\ln^3 d}{dp}(1+o(1))\Bigg] \\
&= \exp\left[\left(\frac{1}{\ln q} - \frac{1}{p}\right)\ln^2 d - \frac{3\ln d\ln\ln d}{\ln q}(1+o(1))\right],
\end{aligned}$$

where the second inequality follows from Observation 2.1, and in the last equality we use the fact that $\frac{\ln^3 d}{dp} = o\left(\frac{\ln d \ln\ln d}{\ln q}\right)$. Now,

$$\frac{1}{\ln q} - \frac{1}{p} = \frac{1}{p + p^2/2 + \mathcal{O}\left(p^3\right)} - \frac{1}{p} = -0.5 + o(1).$$

Therefore, $\mathbb{E}(X_r) \le \exp\left[(-0.5 + o(1))\ln^2 d\right] < \exp[-\ln d]$ for large n, so $\hat{r} \ge \left\lfloor \log_q\left(\frac{d}{\ln^2 d}\right)\right\rfloor$ by definition.

For the upper bound let us redefine $r = \left\lceil \log_q\left(\frac{d}{\ln^2 d}\left(1 + p + \frac{1}{\ln\ln d}\right)\right)\right\rceil$. Then $r = \log_q\left(\frac{d}{\ln^2 d}\left(1 + o(1)\right)\right)$, so Observation 2.1 applies. Thus, for n being sufficiently large,

$$\begin{aligned}
\mathbb{E}(X_r) &= \binom{n}{r}\left(1 - (1-p)^r\right)^{n-r} \\
&\ge \exp\left[r(\ln(n/r) - (n-r)(1-p)^r - (n-r)(1-p)^{2r}\right] \\
&\ge \exp\Bigg[\left(\frac{\ln d}{\ln q} - \frac{2\ln\ln d}{\ln q}(1+o(1))\right)\cdot \ln\left(\frac{d}{\ln d}(1+o(1))\right) \\
&\qquad - \frac{n\ln^2 d}{d\left(1 + p + \frac{1}{\ln\ln d}\right)} + \frac{\ln^3 d}{dp}(1+o(1)) - \frac{\ln^4 d}{dp}(1+o(1))\Bigg] \\
&= \exp\left[\left(\frac{1}{\ln q} - \frac{1}{p\left(1 + p + \frac{1}{\ln\ln d}\right)}\right)\ln^2 d - \frac{3\ln d\,\ln\ln d}{\ln q}(1+o(1))\right],
\end{aligned}$$

where in the equality we use the fact that $\frac{\ln^4 d}{dp} = o\left(\frac{\ln d \ln\ln d}{\ln q}\right)$. Now,

$$\begin{aligned}
\ln^2 d\left(\frac{1}{\ln q} - \frac{1}{p\left(1 + p + \frac{1}{\ln\ln d}\right)}\right) &= \frac{\ln^2 d}{\ln q}\left(1 - \frac{p + p^2/2 + \mathcal{O}(p^3)}{p\left(1 + p + \frac{1}{\ln\ln d}\right)}\right) \\
&= \frac{\ln^2 d}{\ln q}\left(\frac{p}{2} + \frac{1}{\ln\ln d} + \mathcal{O}(p^2)\right)(1+o(1)) \\
&= \left(\frac{\ln^2 d}{2} + \frac{\ln^2 d}{\ln q\,\ln\ln d}\right)(1+o(1)).
\end{aligned}$$

Therefore,

$$\mathbb{E}(X_r) > \exp\left[(0.5 + o(1))\ln^2 d\right] > \exp[-\ln d]$$

for large n, so by definition, $\hat{r} < \left\lceil \log_q\left(\frac{d}{\ln^2 d}\left(1 + p + \frac{1}{\ln\ln d}\right)\right)\right\rceil$.

Part (i) then follows from the minimality of $\hat{r}$ and since $d \to \infty$.

Part (ii) follows from the definition of $\hat{r}$ and by Lemma 2.2:

$$\mathbb{E}(X_{\hat{r}+2}) = \exp\left[(1 + o(1))\ln^2 d\right]\cdot \mathbb{E}(X_{\hat{r}+1}) \ge \exp\left[(1+o(1))\ln^2 d - \ln d\right] \to \infty.$$

$\square$

Note that from part (i) of the previous lemma it follows by the standard first moment argument that

$$\mathbb{P}(D(G) \le \hat{r}) = \mathbb{P}(X_{\hat{r}} > 0) \le \mathbb{E}(X_{\hat{r}}) \to 0.$$

Hence, $\mathbb{P}(D(G) \geq \hat{r}+1) \to 1$. This proves the lower bound of the interval of concentration in the sparse range of the edge probability p in Theorem 1.1.

Part *(ii)* is of course only the first step in deducing $\mathbb{P}(D(G) \leq \hat{r}+2) \to 1$ for the upper bound of the interval. The full analysis happens in Section 3 via the second moment method.

2.2. **The dense case.** Now we take a look at the behaviour of the expected number of dominating sets of size r when the edge probability p tends to 1 moderately fast.

Lemma 2.5. *Let $p \to 1$ such that $\frac{1}{1-p} = q \leq n$. For $r = \log_q\left(\frac{n \ln q}{\ln^2 n}\right)$ we have*

$$\mathbb{E}(X_{\lfloor r \rfloor}) \to 0.$$

Proof. Note that $r = \frac{\ln n - 2\ln\ln n + \ln\ln q}{\ln q} \leq \frac{\ln n}{\ln q} \leq \ln n$ and $(1-p)^r = \frac{\ln^2 n}{n \ln q} \to 0$ since $q \to \infty$.

$$\begin{aligned}
\mathbb{E}(X_{\lfloor r \rfloor}) &= \binom{n}{\lfloor r \rfloor} \cdot \left(1-(1-p)^{\lfloor r \rfloor}\right)^{n-\lfloor r \rfloor} \\
&\leq \left(\frac{ne}{\lfloor r \rfloor}\right)^{\lfloor r \rfloor} \cdot \exp\left[-(n-\lfloor r \rfloor)(1-p)^{\lfloor r \rfloor}\right] \\
&\leq \exp\left[r\Big(\ln n + 1 - \ln r\Big) - (n-r)(1-p)^r\right] \\
&= \exp\left[(-2\ln\ln n + \ln\ln q)\frac{\ln n}{\ln q} + r((1-p)^r + 1 - \ln r)\right] \to 0,
\end{aligned}$$

since $q \leq n$ and $r((1-p)^r + 1 - \ln r) \leq r(2 - \ln r)$ is bounded from above by the constant e. $\square$

One can also show that, analogously to the sparse case, $\mathbb{E}(X_{\lfloor r \rfloor + 2}) \to \infty$. Since we don't need this fact further, we omit the calculation.

3. The variance

In this section, we prove Theorem 1.1. Since the validity of the theorem was shown already for constant p in [**?**], we restrict our attention to the cases when p tends to 0 or 1. We will refer to three cases

- *the sparse case*, when $p \to 0$ but $p \gg \frac{\ln^2 n}{\sqrt{n}}$. In this case recall that $\hat{r} = \min\{\, r \mid \mathbb{E}(X_r) \geq \exp\left[-\ln d\right]\} - 1$ and let us set $r = \hat{r}+2$;
- *the dense case*, when $p \to 1$ but $q = \frac{1}{1-p} \leq n$. In this case we set $\hat{r} = \log_q\left(\frac{n \ln q}{\ln^2 n}\right)$ and $r = \lfloor \hat{r} \rfloor + 2$;
- *the very dense case*, when $q = \frac{1}{1-p} > n$.

The third case is straightforward and will be treated separately at the end of this section.

The calculations for the sparse and the dense case are often identical, so we treat these cases in parallel. We want to apply Chebyshev's Inequality and conclude that

$$\mathbb{P}(X_r = 0) \leq \frac{\mathbb{V}(X_r)}{\mathbb{E}(X_r)^2} \to 0.$$

In the proof we will stumble over some expressions more than once, so we bundle the information of their asymptotic behaviour in the next observation.

Observation 3.1. *In both, the sparse and the dense case, we have,*

(*i*) $(1-p)^r \to 0$,
(*ii*) $r^2 = o(n)$,
(*iii*) $r(1-p)^r \to 0$,
(*iv*) $n(1-p)^{2r-1} \to 0$.

Proof. We treat the two cases separately.
The sparse case. By Lemma 2.4, $\hat{r}$ and thus $r = \hat{r}+2$ are of the form $\log_q\left(\frac{d}{\ln^2 d}(1+o(1))\right)$, and thus Observation 2.1 applies. Part (*i*) is just Observation 2.1 (*ii*). For the rest we use Observation 2.1 and that $p \gg \frac{\ln^2 d}{\sqrt{n}}$ and $d \to \infty$:

$$\frac{r^2}{n} = \frac{\ln^2 d}{np^2}(1+o(1)) \to 0, \qquad r(1-p)^r = \frac{\ln^3 d}{np^2}(1+o(1)) \to 0, \qquad \text{and} \qquad n(1-p)^{2r} = \frac{\ln^4 d}{np^2}(1+o(1)) \to 0.$$

The dense case. By definition, $r = \left\lfloor \frac{\ln n - 2\ln\ln n + \ln\ln q}{\ln q} \right\rfloor + 2$. Therefore, for large n,

$$\log_q\left(\frac{n \ln q}{\ln^2 n}\right) + 1 \le r \le \frac{\ln n}{\ln q} + 2,$$

since $q \le n$. Hence $r \ge 1$ and part (*i*) follows since $p \to 1$. For part (*ii*) we have

$$\frac{r^2}{n} \le \frac{1}{n}\left(\frac{\ln n}{\ln q} + 2\right)^2 = o(1).$$

For part (*iii*) and (*iv*) we have

$$r(1-p)^r \le \left(\frac{\ln n}{\ln q} + 2\right)\frac{\ln^2 n}{n \ln q}(1-p) = o(1),$$
$$n(1-p)^{2r-1} \le \frac{\ln^4 n}{n \ln^2 q}(1-p) = o(1).$$

□

For the variance $\mathbb{V}(X_r) = \mathbb{E}(X_r^2) - \mathbb{E}(X_r)^2$ we need to calculate $\mathbb{E}(X_r^2)$.
Let I_A be the indicator random variable of the event that subset $A \subseteq [n]$ is dominating. Then

$$\mathbb{E}(X_r^2) = \sum_{A,B \in \binom{[n]}{r}} \mathbb{E}(I_A \cdot I_B) = \sum_{A \in \binom{[n]}{r}} \sum_{s=0}^{r} \sum_{\substack{B \in \binom{[n]}{r} \\ |A \cap B| = s}} \mathbb{E}(I_A \cdot I_B)$$

Now, for $A, B \in \binom{[n]}{r}$, $|A \cap B| = s$, we have

$$\begin{aligned}
\mathbb{E}(I_A \cdot I_B) &\le \mathbb{P}\Big(A \text{ dominates } \overline{A \cup B} \wedge B \text{ dominates } \overline{A \cup B}\Big) \\
&= \mathbb{P}\Big(\forall x \in \overline{A \cup B} \quad \exists y_1 \in A \cap \Gamma(x) \wedge \exists y_2 \in B \cap \Gamma(x)\Big) \\
&= \Big(1 - 2(1-p)^r + (1-p)^{2r-s}\Big)^{n-2r+s}.
\end{aligned}$$

So,

$$\mathbb{V}(X_r) \le \binom{n}{r} \sum_{s=0}^{r} \binom{r}{s}\binom{n-r}{r-s}\Big(1 - 2(1-p)^r + (1-p)^{2r-s}\Big)^{n-2r+s} - \mathbb{E}(X_r)^2. \tag{3}$$

First we see that the $s = 0$ term of the sum is asymptotically at most $\mathbb{E}(X_r)^2$:

$$\begin{aligned}
&\binom{n}{r}\binom{n-r}{r}\left(1-2(1-p)^r+(1-p)^{2r}\right)^{n-2r}\\
&\quad=\binom{n}{r}^2\Big(1-(1-p)^r\Big)^{2(n-r)}\frac{\binom{n-r}{r}}{\binom{n}{r}}\Big(1-(1-p)^r\Big)^{-2r}\\
&\quad\le \mathbb{E}(X_r)^2\cdot\exp\left[2r(1-p)^r+2r(1-p)^{2r}\right]\\
&\quad=\mathbb{E}(X_r)^2(1+o(1))
\end{aligned}$$

The inequality holds because $(1-p)^r \to 0$ by Observation 3.1 (i). The final conclusion follows from Observation 3.1 (iii).

Now we estimate the remaining terms of the sum (3). It turns out that the term of $s = 1$ dominates the rest: Let

$$f(s)=\binom{r}{s}\binom{n-r}{r-s}\Big(1-2(1-p)^r+(1-p)^{2r-s}\Big)^{n-2r+s}.$$

We have just seen that $\mathbb{V}(X_r) \le \binom{n}{r}\sum_{s=1}^{r} f(s) + o(\mathbb{E}(X_r)^2)$. We will prove that for large n,

$$\binom{n}{r}\sum_{s=1}^{r} f(s) \le 3\binom{n}{r} f(1). \tag{4}$$

First let us show that indeed, (4) implies $\mathbb{V}(X_r) = o(\mathbb{E}(X_r)^2)$.

$$\begin{aligned}
\frac{\binom{n}{r}f(1)}{\mathbb{E}(X_r)^2}&=\frac{r\binom{n-r}{r-1}\Big(1-2(1-p)^r+(1-p)^{2r-1}\Big)^{n-2r+1}}{\binom{n}{r}(1-(1-p)^r)^{2(n-r)}}\\
&=\frac{r^2}{n}\cdot\frac{\binom{n-r}{r-1}}{\binom{n-1}{r-1}}\cdot\left(1+\frac{p(1-p)^{2r-1}}{(1-(1-p)^r)^2}\right)^{n-r}\cdot\left(\frac{1}{1-2(1-p)^r+(1-p)^{2r-1}}\right)^{r-1}\\
&\le\frac{r^2}{n}\exp\left[\Big((n-r)p(1-p)^{2r-1}+(r-1)2(1-p)^r\Big)(1+o(1))\right]\\
&\to 0,
\end{aligned}\tag{5}$$

where in the last inequality we use again that $(1-p)^r \to 0$ by Observation 3.1 (i). The final conclusion follows by Observation 3.1 (ii), (iii) and (iv). Thus, $\mathbb{V}(X_r) = o(\mathbb{E}(X_r)^2)$ follows.

So we only need to show that (4) holds. We consider the expression

$$\frac{f(1)}{f(s)}=\frac{r\binom{n-r}{r-1}\Big(1-2(1-p)^r+(1-p)^{2r-1}\Big)^{n-2r+1}}{\binom{r}{s}\binom{n-r}{r-s}\Big(1-2(1-p)^r+(1-p)^{2r-s}\Big)^{n-2r+s}}.$$

Note first that for every $2 \le s \le r$,

$$\begin{aligned}
\frac{\binom{n-r}{r-1}}{\binom{r}{s}\binom{n-r}{r-s}}&=\frac{(n-2r+s)_{s-1}}{(r-1)_{s-1}}\cdot\frac{s!}{(r)_s}\ge\left(\frac{n-2r}{r}\right)^{s-1}\cdot\left(\frac{1}{r}\right)^s\\
&=\exp\left[(s-1)\ln\left(\frac{n-2r}{r}\right)-s\ln r\right].
\end{aligned}\tag{6}$$

Also, since $(1-p)^r \to 0$ by Observation 3.1 (i) and $r = o(n)$ by Observation 3.1 (ii), for every $2 \le s \le r$ we have that

$$\frac{\left(1-2(1-p)^r+(1-p)^{2r-1}\right)^{n-2r+1}}{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}} \ge \left(1-\frac{(1-p)^{2r-s}-(1-p)^{2r-1}}{1-2(1-p)^r+(1-p)^{2r-s}}\right)^{n-2r+1}$$
$$= \left(1-(1-p)^{2r-s}(1-(1-p)^{s-1})(1+o(1))\right)^{n-2r+1}$$
$$= \exp\Big[-n(1-p)^{2r-s}(1-(1-p)^{s-1})(1+o(1))\Big]. \tag{7}$$

From now on, we need to separate the sparse and the dense case.

3.1. **The sparse case.** Recall that in this case $p \gg \frac{\ln^2 d}{\sqrt{n}}$ and $d \to \infty$. In order to deduce (4), we show for n large enough that for every $2 \le s \le \min\{\ln n, 1/\sqrt{p}\}$, we have $f(1) \ge \ln n\, f(s)$, and for every $\min\{\ln n, 1/\sqrt{p}\} \le s \le r$, we have that $f(1) \ge r f(s)$. Then we have that

$$\sum_{s=1}^{r} f(s) \le f(1) + \ln n \, \max\left\{f(s) : 2 \le s \le \min\left\{\ln n, 1/\sqrt{p}\right\}\right\}$$
$$+ r \max\left\{f(s) : \min\left\{\ln n, 1/\sqrt{p}\right\} \le s \le r\right\}$$
$$\le 3f(1).$$

We split the analysis into three cases.

Small range. First, suppose $2 \le s \le \min\left\{\ln n, 1/\sqrt{p}\right\}$. Then, since $s \ll 1/p$, we have that

$$(1-p)^s \ge 1 - ps \qquad \text{and thus} \qquad (1-p)^s = 1 + o(1). \tag{8}$$

So,

$$\frac{f(1)}{\ln n \cdot f(s)} = \frac{r}{\ln n} \cdot \frac{\binom{n-r}{r-1}}{\binom{r}{s}\binom{n-r}{r-s}} \cdot \frac{\left(1-2(1-p)^r+(1-p)^{2r-1}\right)^{n-2r+1}}{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}}$$
$$\overset{(6),(7),(8)}{\ge} \exp\Bigg[\ln r - \ln\ln n + (s-1)\ln\left(\frac{n-2r}{r}\right) - s\ln r$$
$$- np(s-1)(1-p)^{2r-s}(1+o(1))\Bigg]$$
$$\overset{(8)}{=} \exp\left[(s-1)\left(\ln\left(\frac{n}{r}\right) + o(1) - \ln r - np(1-p)^{2r}(1+o(1))\right) - \ln\ln n\right]$$
$$\ge \exp\left[(s-1)\left((1+o(1))\ln\left(\frac{n}{r^2\ln n}\right) + o(1)\right)\right]$$
$$\ge 1,$$

where the second to last inequality follows from Observation 3.1 (iv), whereas the last one follows since $s \ge 2$, and from $\frac{n}{r^2 \ln n} \ge \frac{np^2}{\ln^3 n}(1+o(1)) \to \infty$ by Observation 2.1 (i) and since

$p \gg \frac{\ln^{3/2} n}{\sqrt{n}}$.

Middle range. Now, let $\min\left\{\ln n, \frac{1}{\sqrt{p}}\right\} \le s \le 0.9\, r$. Since $\min\left\{\ln n, \frac{1}{\sqrt{p}}\right\} \to \infty$, by (6) and (7) we obtain

$$\begin{aligned}
\frac{f(1)}{r\cdot f(s)} &= \frac{\binom{n-r}{r-1}}{\binom{r}{s}\binom{n-r}{r-s}} \cdot \frac{\left(1-2(1-p)^r+(1-p)^{2r-1}\right)^{n-2r+1}}{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}} \\
&\ge \exp\left[s\ln\left(\frac{n}{r^2}\right)(1+o(1)) - n(1-p)^{2r-s}(1+o(1))\right]. \qquad (9)
\end{aligned}$$

We will show that $s\,\ln n\left(\frac{n}{r^2}\right) \gg n(1-p)^{2r-s}$ and conclude that $f(1) \ge r\, f(s)$. First note that $\ln\left(\frac{n}{r^2}\right) \to \infty$ by Observation 3.1 (*ii*). Also, by Observation 2.1 $n(1-p)^{2r} = \frac{\ln^4 d}{np^2}(1+o(1))$. By differentiating one finds that the function $g(s) := s(1-p)^s$ takes its minima at the endpoints of the interval $\left[\min\left\{\ln n, \frac{1}{\sqrt{p}}\right\}, 0.9r\right]$. We check that both values $g\left(\min\left\{\ln n, \frac{1}{\sqrt{p}}\right\}\right)$ and $g(0.9\,r)$ have higher order than $\frac{\ln^4 d}{np^2}$, and hence for large enough n the exponent of (9) is positive, completing the proof of the middle range. Firstly by Observation 2.1,

$$0.9\,r(1-p)^{0.9\,r} = 0.9\,\frac{\ln d}{p}\left(\frac{\ln^2 d}{d}\right)^{0.9}(1+o(1)) \gg \frac{\ln^4 d}{np^2}.$$

Secondly since $p \to 0$ and by Observation 3.1 (*iv*),

$$\frac{1}{\sqrt{p}}(1-p)^{\frac{1}{\sqrt{p}}} = \frac{1}{\sqrt{p}}(1-o(1)) \gg 1 \gg n(1-p)^{2r}.$$

We need to bound $g(\ln n)$ only if $p < 1/\ln^2 n$, otherwise $\min\left\{\ln n, \frac{1}{\sqrt{p}}\right\} = \frac{1}{\sqrt{p}}$. But then

$$\ln n(1-p)^{\ln n} = \ln n \exp(-p\ln n(1+o(1))) \ge \ln n\exp(-1/\ln n)(1+o(1)) \gg n(1-p)^{2r}.$$

This completes the proof of the middle range.

Large range. Finally, let $0.9r \le s \le r$. Then

$$\frac{f(1)}{r\,f(s)} = \frac{\binom{n-r}{r-1}}{\binom{r}{s}\binom{n-r}{r-s}} \cdot \frac{\left(1-2(1-p)^r+(1-p)^{2r-1}\right)^{n-2r+1}}{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{n-2r+s}}$$

$$= \mathbb{E}(X_r)\cdot \frac{\binom{n-r}{r-1}}{\binom{n}{r}\binom{r}{s}\binom{n-r}{r-s}} \cdot \left[\frac{1-2(1-p)^r+(1-p)^{2r-1}}{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)(1-(1-p)^r)}\right]^{n-r}$$

$$\cdot \frac{\left(1-2(1-p)^r+(1-p)^{2r-s}\right)^{r-s}}{\left(1-2(1-p)^r+(1-p)^{2r-1}\right)^{r-1}}$$

$$\ge \mathbb{E}(X_r)\frac{r(1+o(1))}{n}\cdot\frac{1}{\binom{r}{s}\binom{n-r}{r-s}}\cdot\left[1+\frac{(1-p)^r(1-(1-p)^{r-s})}{1-2(1-p)^r+(1-p)^{2r-s}}\right]^{n-r}$$

$$\ge \exp\Big[(1+o(1))\,\ln^2 d - \ln n - 2(r-s)\,\ln n + n(1-p)^r(1-(1-p)^{r-s})(1+o(1))\Big], \tag{10}$$

where in the first inequality, we estimated the last factor by 1 and used that

$$\frac{\binom{n-r}{r-1}}{\binom{n}{r}} \ge \frac{r}{n}\left(1-\frac{r-1}{n-r+1}\right)^{r-1} = \frac{r}{n}(1+o(1)),$$

since $r^2 = o(n)$ by Observation 3.1 (ii). In the last inequality we estimated by $\binom{r}{s} = \binom{r}{r-s} \le n^{r-s}$ and $\binom{n-r}{r-s} \le n^{r-s}$ as well as used that $r = o(n)$ and $(1-p)^r = o(1)$, by Observation 2.1 (i) and (ii), and Lemma 2.4 (ii). Now since $d \gg \sqrt{n}\ln^2 n$, $(1+o(1))\,\ln^2 d - \ln n \ge 1/4\ln^2 n(1+o(1))$ in (10). Note that the $1+o(1)$-function in (10) does not depend on s. So for sufficiently large n, this expression is larger than $1/2$, and thus, it is enough to show that for large n and every $s \in [0.9r, r]$ in the large range

$$-2(r-s)\,\ln n + \frac{1}{2}n(1-p)^r(1-(1-p)^{r-s}) \ge 0. \tag{11}$$

To prove (11) set $x = r-s$ and consider the function $h(x) := -2x\,\ln n + \frac{1}{2}n(1-p)^r(1-(1-p)^x)$ on the interval $[0, 0.1r]$. By differentiating twice we see that $h(x)$ is concave, and therefore $h(x) \ge \min\{h(0), h(0.1\,r)\}$ for $0 \le x \le 0.1r$. Now $h(0) = 0$. For the other endpoint, by Observation 2.1 (i) and (ii), and since $d \gg \sqrt{n}$ we have

$$h(0.1r) = -0.2r\ln n + \frac{1}{2}n(1-p)^r(1-(1-p)^{0.1\,r})$$
$$= -0.2r\ln n + \frac{r\ln d}{2}(1-o(1)) \to \infty.$$

This finishes the proof of the large range, and therefore also the proof of Theorem 1.1 in the sparse case.

3.2. **The dense case.** Similarly to the sparse case, we aim to prove $\sum_{s=1}^r f(s) \le 3f(1)$. We show in the following that for every $2 \le s \le r$ we have $f(1) \ge r\,f(s)$. This then implies (4). Recall that $q = \frac{1}{1-p} \to \infty$, $q \le n$ and that $r = \left\lfloor \log_q\left(\frac{n\ln q}{\ln^2 n}\right)\right\rfloor + 2$, that is

$$\frac{\ln n + \ln\ln q - 2\ln\ln n}{\ln q} + 1 \le r \le \frac{\ln n + \ln\ln q - 2\ln\ln n}{\ln q} + 2.$$

Let $2 \le s \le r$. First note that $(1-p)^{s-1} \to 0$ since $p \to 1$. Then by the inequalities (6) and (7),

$$\begin{aligned}
\frac{f(1)}{r\,f(s)} &\geq \exp\left[(s-1)\ln\left(\frac{n-2r}{r}\right) - s\ln r - n(1-p)^{2r-s}(1-(1-p)^{s-1})(1+o(1))\right]\\
&\geq \exp\left[(s-1)\Bigg(\ln n + o(1) - 2\left(\frac{s}{s-1}\right)\ln r\Bigg) - n(1-p)^{2r-s}(1+o(1))\right]\\
&= \exp\left[(s-1)\ln n(1+o(1)) - n(1-p)^{2r-s}(1+o(1))\right],
\end{aligned}$$

since $\ln r \le \ln\left(\frac{\ln n}{\ln q}\right) \le \ln\ln n \ll \ln n$ and $\frac{s}{s-1} \le 2$. We will show that $(s-1)\ln n \ge 2n(1-p)^{2r-s}$ for all $2 \le s \le r$. To this end, consider the function $h(x) := x\ln n - 2n(1-p)^{2r-1}(1-p)^{-x}$ on the interval $[1, r-1]$. Differentiating twice shows that h is concave, and thus, $h(x) \ge \min\{h(1), h(r-1)\}$ for $1 \le x \le r-1$. Now, $h(1) = \ln n - 2n(1-p)^{2r-2} = \ln n - o(1) \gg 1$ by Observation 3.1 *(iii)*. Secondly,

$$\begin{aligned}
h(r-1) &= (r-1)\ln n - 2n(1-p)^r\\
&\geq \frac{\ln n + \ln\ln q - 2\ln\ln n}{\ln q}\ln n - 2\frac{\ln^2 n}{q\,\ln q}\\
&\geq \frac{\ln^2 n}{\ln q}(1-o(1)) \geq \ln n \gg 1,
\end{aligned}$$

where we used that $(1-p)^r \le \frac{\ln^2 n}{n\ln q}(1-p)$, $q \to \infty$ and that $q \le n$. We conclude that $f(1) \gg r f(s)$ for all $2 \le s \le r$. This finishes the proof of the dense case.

3.3. **The very dense case.** Now, we complete the picture when p tends to 1.
Let $p = p(n) \to 1$ such that $q > n$. Then, $\log_q\left(\frac{n\ln q}{\ln^2 n}\right) < 1$ since $\frac{n}{\ln^2 n} < \frac{q}{\ln q}$ for $n \ge 3$. Thus $r = \lfloor \hat{r} \rfloor + 2 \le 2$ and we claim that the domination number is a.a.s. at most 2. In fact, when $p \ge 1 - 1/n$, then the complement of $G \sim \mathcal{G}(n,p)$ has an isolated vertex a.a.s. But the very same vertex is a dominating set of size 1 in G.

The proof of Theorem 1.1 is complete.

4. Concentration for smaller values of p

In this section we prove Proposition 1.2 and Theorem 1.3.

Proof. (of Proposition 1.2) Let $\hat{r}$ be given by (1). Then by Lemma 2.4 and Observation 2.1, $\hat{r} = \log_q\left(\frac{d}{\ln^2 d}(1+o(1))\right) = \frac{n}{d}\ln d\,(1+o(1))$ and $\mathbb{E}(X_{\hat{r}}) \to 0$. Therefore,

$$\mathbb{P}(D(G) \le \hat{r}) = \mathbb{P}(X_{\hat{r}} > 0) \le \mathbb{E}(X_{\hat{r}}) \to 0,$$

and hence, $D(\mathcal{G}(n,p)) > \hat{r}$ a.a.s., proving the lower bound.

For the upper bound, we set $r = n\frac{\ln d}{d}$ and apply the alteration technique from [**?**, Theorem 1.2.2]. We take the set $[r]$ and calculate how many vertices are *not* dominated by it. Adding these vertices to $[r]$ results in a dominating set. Let Y be the number of vertices not dominated by $[r]$. Then Y is the sum of $n-r$ independent Bernoulli trials with success probability $(1-p)^r$

each. By Markov's Inequality, $\mathbb{P}(Y \leq \mathbb{E}(Y) \ln\ln d) \to 0$ since $d \to \infty$. So with probability tending to 1, $\mathcal{G}(n,p)$ has a dominating set of size

$$r+\mathbb{E}(Y)\ln\ln d = r+(n-r)(1-p)^r \ln\ln d \leq n\frac{\ln d}{d}+n\exp\left(-pn\frac{\ln d}{d}\right)\ln\ln d = n\frac{\ln d}{d}(1+o(1)).$$

Therefore, $D(\mathcal{G}(n,p)) \leq n\frac{\ln d}{d}(1+o(1))$ a.a.s. □

We now prove Theorem 1.3.

Proof of Theorem 1.3. We use Talagrand's Inequality to show concentration in Theorem 1.3. To that end, let us introduce the necessary terminology. The following setting can be found in [**?**, Chapter 7.7].
Let $\Omega = \prod_{i=1}^{N} \Omega_i$ be the product space of probability spaces Ω_i, equipped with the product measure. We say that a random variable $X : \Omega \to \mathbb{R}$ is *Lipschitz*, if $|X(x) - X(y)| \leq 1$ whenever x and y differ in at most one coordinate. Further, for a function $f : \mathbb{N} \to \mathbb{N}$, we say that X is *f-certifiable* if whenever $X(x) \geq s$ there exists $I \subseteq [N]$ with $|I| \leq f(s)$ such that for all $y \in \Omega$ with $x_I = y_I$ we have $X(y) \geq s$. We use the following version of Talagrand's Inequality.

Theorem 4.1 (Talagrand's Inequality)**.** *Let $f : \mathbb{N} \to \mathbb{N}$ be a function, and suppose $X : \Omega \to \mathbb{R}$ is a random variable that is Lipschitz and f-certifiable. Then for all a, $u \in \mathbb{R}$:*

$$\mathbb{P}\Big(X \leq a - u\sqrt{f(a)}\Big) \cdot \mathbb{P}(X \geq a) \leq e^{-u^2/4}.$$

Corollary 4.2. *For all $b, t \in \mathbb{R}$,*

$$\mathbb{P}(D(G) \leq b) \cdot \mathbb{P}(D(G) \geq b+t) \leq e^{-t^2/4(n-b)}. \tag{12}$$

Proof. We check that Theorem 4.1 can be applied to our situation. For that reason, we identify a graph G with its edge set, and view $\mathcal{G}(n,p)$ as the product of $N = \binom{n}{2}$ Bernoulli experiments with parameter p. Let us consider the random variable $X : \mathcal{G}(n,p) \to \mathbb{R}$ defined by $X(G) = n - D(G)$. Clearly, X is Lipschitz, since adding or deleting an edge changes the domination number (and hence X) by at most one. Further, X is f-certifiable, where $f(s) = s$. To see this, assume $X(G) \geq s$, i.e. $D(G) \leq n - s$. Then there exists a dominating set S of size $n - s$. We can choose s edges, one from each $v \in (V(G) \setminus S)$ to S, which certify that $D(G) \leq n - s$ (more precisely that S is a dominating set). Clearly, any graph H that contains those s edges will have $D(H) \leq n - s$, or $X(H) \geq s$, respectively. Now, it follows by Theorem 4.1, that for all $a, u \in \mathbb{R}$,

$$\mathbb{P}\Big(n - D(G) \leq a - u\sqrt{a}\Big) \cdot \mathbb{P}\Big(n - D(G) \geq a\Big) \leq e^{-u^2/4}.$$

Substituting $b = n - a$ and $t = u\sqrt{a}$ proves the claim. □

To turn Corollary 4.2 into a meaningful result let $t = t(n)$ be any function such that $t = \omega(\sqrt{n})$. If we now set b to be the median in Corollary 4.2, then we obtain $\mathbb{P}(D(G) \geq m+t) \leq 2e^{-t^2/4n}$. Analogously, setting $b + t = m$ gives $\mathbb{P}(D(G) \leq m - t) \leq 2e^{-t^2/4n}$. Hence, Theorem 1.3 follows. □

5. Non-concentration

In this section we prove Theorem 1.4, giving a justification for the existence of a lower bound on p in Theorem 1.1.

Let us first give an outline of the proof. We assume to the contrary that for some $C = c\frac{\hat{r}}{n\sqrt{p}}$, we have that $D(G) \leq \hat{r} + C$ a.a.s for $G \sim \mathcal{G}(n,p)$. Then we delete edges of G with a tiny probability, a probability so small, that the resulting graph $F \sim \mathcal{G}(n,p')$ (where p' is very close to p) shows very similar properties to G. In particular, it will be true that $D(F) \leq \hat{r}+C$ still holds a.a.s. On the other hand we will also show that the deletion process ruins every single dominating set of size $\hat{r}+C$ with positive probability, a contradiction.

A graph property (set of graphs) Q is called *convex*, if for any three graphs $G \subseteq F \subseteq H$, from $G \in Q$ and $H \in Q$ one obtains $F \in Q$. Since the graph property defined by the domination number being at most $\hat{r}+C$ is monotone increasing, it is also convex and the following proposition can be applied.

Proposition 5.1. *Let Q be a convex graph property, $p(1-p)\binom{n}{2} \to \infty$, $x \in \mathbb{R}$, and set $p' = p + x\frac{\sqrt{p}}{n}$. Further, suppose that $G \sim \mathcal{G}(n,p)$ and $F \sim \mathcal{G}(n,p')$. Then*

$$G \in Q \text{ a.a.s.} \;\Rightarrow F \in Q \text{ a.a.s.}$$

Proof. This follows easily from Theorem 2.2 (ii) in [**?**]: if Q is a convex graph property and $p(1-p)\binom{n}{2} \to \infty$, then $\mathcal{G}(n,p)$ has property Q a.a.s. if and only if for all fixed X the graph $\mathcal{G}(n,M)$ has Q a.a.s., where $M = \left\lfloor p\binom{n}{2} + X\sqrt{p(1-p)\binom{n}{2}}\right\rfloor$. □

We need one more definition for the proof. Let $G = (V,E)$ be a graph. For a subset $S \subseteq V$ of the vertices, we call an edge $e = xs \in E$ *crucial w.r.t.* S if $s \in S$, $x \in V \setminus S$ and for all $s' \in S \setminus \{s\}$, $xs' \notin E$. That is, a crucial edge is the only connection of x into S in G. In particular, if S was dominating, the deletion of e would result in S not being dominating anymore. Set

$$\mathcal{C}_G(S) = \{x \in V \setminus S |\, xs \in E(G) \text{ is crucial w.r.t. } S \text{ for some } s \in S\}.$$

Note that by the definition of a crucial edge, $|\mathcal{C}_G(S)|$ counts exactly the number of crucial edges. We are now ready to prove the main theorem of this section.

Proof of Theorem 1.4. Suppose to the contrary that there exist $\varepsilon, c > 0$, such that for every $\delta = 1/k$, there exists $p_k = p_k(n)$ with $\varepsilon/n \leq p_k \ll 1$ and an increasing sequence $\left(\pi^{(k)}(n)\right)_{n\in\mathbb{N}}$ of positive integers satisfying

$$\mathbb{P}\left(D\left(\mathcal{G}\left(\pi^{(k)}(n), p_k\left(\pi^{(k)}(n)\right)\right)\right) > \hat{r}\left(\pi^{(k)}(n)\right) + c\frac{\hat{r}\left(\pi^{(k)}(n)\right)}{\pi^{(k)}(n)\sqrt{p_k\left(\pi^{(k)}(n)\right)}}\right) < 1/k.$$

From these probabilities p_k and sequences $\pi^{(k)}(n)$ we can define $p = p(n)$ with $1/\varepsilon \leq p \ll 1$ and a new increasing sequence $(\pi(n))_{n\in\mathbb{N}}$ of positive integers such that

$$\mathbb{P}\left(D(\mathcal{G}(\pi(n), p(\pi(n)))) > \hat{r}(\pi(n)) + c\frac{\hat{r}(\pi(n))}{\pi(n)\sqrt{p(\pi(n))}}\right) < 1/n \to 0.$$

The sequence $(\pi(n))_{n\in\mathbb{N}}$ than has a subsequence $(\tau(n))_{n\in\mathbb{N}}$ such that either $p(\tau(n)) \gg 1/n$, or there exists a constant $K > 0$ with $p(\tau(n)) \leq K/n$ for every n. We will deal with both cases simultaneously, splitting the proof into a short case distinction whenever necessary and

reaching a contradiction at the end. For simplicity of notation, we assume that τ is the identity function, as the proof obviously follows the same lines whenever we restrict to a subsequence of the natural numbers.

Recall that according to (1) and Lemma 2.4, for $p \gg 1/n$ we have

$$\hat{r} = \min \{ r \mid \mathbb{E}(X_r) \geq \exp[-\ln d] \} - 1 = \log_q \left(\frac{d}{\ln^2 d} (1 + o(1)) \right).$$

We now show that if $\varepsilon/n \leq p \leq K/n$ then $\hat{r} = \Omega(n)$ as well as $\hat{r} = (1 - \Omega(1))n$. (Here and later we write Ω and Θ for $\Omega_{\varepsilon,K,c}$ and $\Theta_{\varepsilon,K,c}$, respectively.)

If $\hat{r} = o(n)$, then we see that

$$\mathbb{E}(X_{\hat{r}+1}) = \binom{n}{\hat{r}+1} \left(1 - (1-p)^{\hat{r}+1}\right)^{n-\hat{r}-1} = 2^{o(n)} (o(1))^{n-o(n)} = o(1),$$

a contradiction to the definition of $\hat{r}$, since $1/d = \Omega(1)$. So $\hat{r} = \Omega(n)$. On the other hand, the expected number of vertices of degree exactly 1 in G is at least $\varepsilon(n-1)(1-K/n)^{n-1}$ and hence (say, by Chebyshev's inequality) a.a.s. there are at least $\frac{\varepsilon}{2} e^{-2K} n$ vertices of degree exactly 1 in G. Since deleting all isolated vertices and all vertices of degree at least 2 leaves us with a graph of maximum degree one, we obtain $D(G) \leq \left(1 - \frac{\varepsilon}{4} e^{-2K}\right) n$ a.a.s. Every set containing a dominating set is itself dominating, so there are a.a.s. at least $\frac{\varepsilon}{4} e^{-2K} n$ dominating sets of size $r^* = n\left(1 - \frac{\varepsilon}{4} e^{-2K}\right) + 1$. Therefore $\mathbb{E}(X_{r^*}) \geq \frac{\varepsilon}{4} e^{-2K} n(1 - o(1)) \geq 1/d$ for large enough n and we obtain that $\hat{r} \leq r^* - 1 = n(1 - \Theta(1))$.

We denote $C = \left\lfloor c \frac{\hat{r}}{n\sqrt{p}} \right\rfloor$, $r = \hat{r} + C$, and $I = \{1, 2, \ldots, r\}$. Note that $\hat{r} = \Theta(n)$ implies $C = \Theta(\sqrt{n})$.

Since we are aiming to use Lemma 2.2 about the expected value of the number of dominating sets of size $\hat{r} + C$, we first show that a similar statement holds when $\varepsilon/n \leq p \leq K/n$.

$$\begin{aligned}
\mathbb{E}(X_{\hat{r}+C}) &< \frac{\mathbb{E}(X_{\hat{r}+C})}{\mathbb{E}(X_{\hat{r}})} \cdot \frac{1}{d} \\
&= \frac{\binom{n}{\hat{r}+C}}{\binom{n}{\hat{r}}} \cdot \left(\frac{1 - (1-p)^{\hat{r}+C}}{1 - (1-p)^{\hat{r}}} \right)^{n-\hat{r}-C} \cdot \frac{1}{\left(1 - (1-p)^{\hat{r}}\right)^C} \cdot \frac{1}{d} \\
&< \left(\frac{n - \hat{r}}{\hat{r}} \right)^C \cdot \left(1 + \frac{(1-p)^{\hat{r}} \left(1 - (1-p)^C\right)}{1 - (1-p)^{\hat{r}}} \right)^{n-\hat{r}-C} \cdot \exp\left(\Theta\left(\sqrt{n}\right)\right) \\
&= \exp\left(\Theta\left(\sqrt{n}\right)\right) \left(1 + \Theta\left(1/\sqrt{n}\right)\right)^{\Theta(n)} = \exp\left(\Theta\left(\sqrt{n}\right)\right).
\end{aligned} \tag{13}$$

Similarly we obtain $\mathbb{E}(X_{\hat{r}-C}) = \exp(-\Theta(\sqrt{n})) = o(1)$, hence a.a.s. $D(G) > \hat{r} - C$.

For $p \gg 1/n$, we see that $C = o(1/p)$. Therefore, Observation 2.1 and Lemma 2.2 apply for r.

Let now $p > \varepsilon/n$ be again arbitrary with $p = o(1)$. Suppose for a contradiction that $D(G) \in I$ a.a.s. Consider the following two-stage random procedure. First, we draw G from $\mathcal{G}(n,p)$. Then from G, we delete every edge with probability $p'' := x(\sqrt{p}n)^{-1}$, these choices being independent with x being a sufficiently large constant depending on ε, c and K (in case $p \leq K/n$); we will specify x later. (Notice that, whenever we use the asymptotic notation Θ or similar, we mean constants that are independent of x.) In the new graph F every edge occurs with probability $p' = p(1 - p'') = p - x\frac{\sqrt{p}}{n}$, hence $F \sim \mathcal{G}(n, p')$. By Proposition 5.1,

we know that

$$D(G) \in I \text{ a.a.s. } \Rightarrow D(F) \in I \text{ a.a.s.} \tag{14}$$

In the following we will show that $D(F) \notin I$ a.a.s. This contradiction completes the proof.

We denote by B the event that there exists a dominating set S of size r in G, such that none of its crucial edges have been destroyed, i.e., $\mathcal{C}_G(S) \subseteq N_F(S)$. Clearly, $\bar{B}$ implies $D(F) \notin I$, since no dominating set of size r in G remains dominating in F (and hence there is no smaller dominating set either). For a subset $S \in \binom{V}{r}$ of the vertices, let Y_S be the random variable counting the deleted crucial edges w.r.t. S and denote by D_S the event that S is dominating in G. By the union bound we have for every $f = f(n) > 0$ that

$$\begin{aligned}
\mathbb{P}(B) &= \mathbb{P}\left(\exists S \in \binom{[n]}{r} : \; D_S \text{ holds and } Y_S = 0\right) \\
&\le \sum_{S \in \binom{[n]}{r}} \mathbb{P}\left(D_S \text{ holds, } Y_S = 0 \text{ and } \; |\mathcal{C}_G(S)| \ge f\right) + \\
&\qquad + \sum_{S \in \binom{[n]}{r}} \mathbb{P}\left(D_S \text{ holds, } \; Y_S = 0 \text{ and } \; |\mathcal{C}_G(S)| < f\right) \\
&\le \sum_{S \in \binom{[n]}{r}} \mathbb{P}\left(D_S\right) \cdot \mathbb{P}\left(Y_S = 0 \mid D_S \text{ holds and } |\mathcal{C}_G(S)| \ge f\right) \\
&\qquad + \sum_{S \in \binom{[n]}{r}} \mathbb{P}\left(D_S\right) \cdot \mathbb{P}\left(|\mathcal{C}_G(S)| < f \mid D_S\right).
\end{aligned} \tag{15}$$

We start by estimating the second sum. Let $S \in \binom{[n]}{r}$. We will observe that $|\mathcal{C}_G(S)|$, conditioned on S being dominating in G, is a binomially distributed random variable. To this end, for vertices $v \in V \setminus S$, define the events

$$A_v = \{v \in \mathcal{C}_G(S) \,|\, D_S\} = \{b_v = 1 \,|\, b_w \neq 0 \; \forall \; w \in V \setminus S\},$$

where b_v is the number of edges of G among the pairs $\{vs : s \in S\}$. The edge sets $\{vs : s \in S\}$ are pairwise disjoint, hence the random variables b_v, and in turn the events A_v are mutually independent.

The random variable $|\mathcal{C}_G(S)|$ conditioned on D_S is then the sum of $(n-r)$ mutually independent random variables: the indicator random variables of the events A_v. These are 1 with probability $p^* := \mathbb{P}(A_v) = \mathbb{P}(b_v = 1)/\mathbb{P}(b_v \neq 0) \ge \mathbb{P}(b_v = 1) = rp(1-p)^{r-1}$. Hence for the expectation of $|\mathcal{C}_G(S)|$ we have

$$\mu := \mathbb{E}(|\mathcal{C}_G(S)| \,|\, D_S) = (n-r)p^* \ge (n-r)pr(1-p)^{r-1}.$$

Thus by Chernoff's inequality (see e.g. [?]), plugging in $f = \mu/2$ we have that

$$\mathbb{P}\left(|\mathcal{C}_G(S)| < \frac{\mu}{2} \;\middle|\; D_S\right) < \exp\left[-\frac{\mu}{8}\right] \tag{16}$$

for large enough n.

Now, we bound the first sum in (15). Observe that conditioning on $|\mathcal{C}_G(S)|$ being a fixed integer ℓ, we have $Y_S \sim \mathrm{Bin}(\ell, p'')$, where p'' is the probability that an edge is deleted from G. Furthermore, observe that once we condition on $|\mathcal{C}_G(S)|$ taking one fixed value, no other information about G influences the distribution of Y_S, especially not the fact that S is dominating

in G, so

$$\mathbb{P}\left(Y_S = 0 \,|\, |\mathcal{C}_G(S)| = \ell \text{ and } D_S\right) = \mathbb{P}\left(Y_S = 0 \,|\, |\mathcal{C}_G(S)| = \ell\right) = (1-p'')^{\ell}.$$

Hence, if G was drawn such that $|\mathcal{C}_G(S)| \geq \mu/2$, then

$$\mathbb{P}\left(Y_S = 0 | D_S \text{ holds and } \ |\mathcal{C}_G(S)| \geq \mu/2\right) \leq (1-p'')^{\mu/2}. \tag{17}$$

Combining (16) and (17) we obtain in (15) that

$$\begin{aligned}\mathbb{P}(B) &\leq \sum_{S \in \binom{[n]}{r}} \mathbb{P}\left(D_S\right) \cdot \Big((1-p'')^{\mu/2} + \exp\left[-\mu/8\right]\Big) \\ &\leq E(X_r) \cdot \exp\left[-p''\mu/3\right], \qquad (18)\end{aligned}$$

since $p'' \to 0$.

We do a case analysis. For $\varepsilon/n \leq p \leq K/n$, recall that $r = \hat{r} + C = \Theta(n)$ and therefore

$$\mu \geq (n-r)pr(1-p)^{r-1} = \Omega(n),$$

and plugging it together with (13) into (18), we obtain

$$\mathbb{P}(B) \leq \exp\left(\Theta\left(\sqrt{n}\right)\right) \cdot \exp\left(-x \cdot \Theta\left(\sqrt{n}\right)\right) \to 0$$

for x sufficiently large depending on ε, c, K.

On the other hand, if $p \gg 1/n$, we see that

$$\mu \geq (n-r)pr(1-p)^{r-1} = (1-o(1))r\ln^2 d$$

by Observation 2.1 (ii). Hence, plugging it into (18) and using Lemma (2.2), we obtain

$$\mathbb{P}(B) \leq \exp\left(C\ln^2 d(1+o(1)) - (1+o(1))x\frac{r\ln^2 d}{3n\sqrt{p}}\right) \to 0$$

for $x \geq 4c$.

Hence, $\mathbb{P}(D(F) \notin I) \geq \mathbb{P}\left(\bar{B}\right) \to 1$ in both cases, a contradiction. $\square$

6. Concluding remarks and open problems

As we already remarked in the introduction, for the range $p = o(1/n)$, we can derive that the domination number of $G \sim \mathcal{G}(n,p)$ is not a.a.s. concentrated on any interval of length $o\left(n\sqrt{p}\right)$. To see this, we can draw G and then delete edges from it with probability $p'' := \left(n\sqrt{p}\right)^{-1}$ to obtain F as described in the previous section. However, in this range the structure of a typical graph G is such that almost all its edges are isolated. Hence, deleting a random edge from G a.a.s. increases the domination number by one. Thus, the domination number of F is a.a.s. $\Theta\left(n\sqrt{p}\right)$ larger than the domination number of G, providing the non-concentration statement.

As it was noted in the introduction, Theorem 1.4 implies that for $p \leq (\ln n/n)^{2/3}$ the domination number $D(\mathcal{G}(n,p))$ is not concentrated on any constant length interval around $\hat{r}$. It would be interesting to improve this result in a couple of directions. On the one hand, it seems reasonable to believe that the power $-\frac{2}{3}$ in the upper bound on the edge probability p could be pushed up to $-\frac{1}{2}$, hence making Theorem 1.1 tight up to a polylogarithmic factor. On the other hand it is unsatisfactory that our current proof of Theorem 1.4 requires the extra assumption that the concentration interval is around $\hat{r}$. It would be desirable to obtain a result stating the non-concentration of $D(\mathcal{G}(n,p))$ on any constant-length interval, independent of its location — like we have it for $p = o(1/n)$.

It would be interesting to learn more about the concentration of the domination number of $\mathcal{G}(n,p)$ in case $p = \mathcal{O}\left(\ln^2 n/\sqrt{n}\right)$. Theorem 1.3 does provide a concentration interval of length slightly above $\sqrt{n}$ for all p. Calculating the $o(1)$-term in Proposition 1.2 gives a bound of the order $n\frac{\ln\ln d}{d}$ on the length of the concentration interval. This is better than the one from Theorem 1.3 for $p \geq \frac{\ln\ln n}{\sqrt{n}}$.

Ideally one would like to know a tight concentration result for every p. Note that Theorems 1.3 and 1.4 imply such a result when $p = \Theta(1/n)$, provided $\hat{r}$ and m are close to each other. In fact, in this range Theorem 1.4 implies that for every interval of length $\mathcal{O}(\sqrt{n})$ around $\hat{r}$, the domination number is not in this interval with some constant positive probability. On the other hand, Theorem 1.3 implies that for every $t \gg \sqrt{n}$, the domination number is a.a.s. concentrated on an interval of length t around its median. That is, under the assumption that $|m - \hat{r}| = \mathcal{O}(\sqrt{n})$, the domination number is not a.a.s. concentrated on any interval of length $\mathcal{O}(\sqrt{n})$, but for every $t \gg \sqrt{n}$, the domination number is a.a.s. concentrated on some interval of length t. Otherwise, if the median m of $D(G)$ is much farther than $\sqrt{n}$ away from $\hat{r}$, then we cannot exclude the possibility that concentration on a short interval has positive probability. However, we would find this outcome quite surprising.

Acknowledgement. We are grateful to Michael Krivelevich for fruitful discussions and especially for his suggestions to the proof of Theorem 1.4.

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany

E-mail address: glebov@mi.fu-berlin.de

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany

E-mail address: liebenau@mi.fu-berlin.de

Institut für Mathematik, Freie Universität Berlin, Arnimallee 3-5, D-14195 Berlin, Germany

E-mail address: szabo@mi.fu-berlin.de